The Rate Region of the Quadratic Gaussian Two-Terminal Source-Coding Problem

Aaron B. Wagner*, Saurabha Tavildar†, and Pramod Viswanath†
*Coordinated Science Laboratory, University of Illinois at Urbana-Champaign and
School of Electrical and Computer Engineering, Cornell University. Email: wagner@ece.cornell.edu.
†Department of Electrical and Computer Engineering and Coordinated Science Laboratory,
University of Illinois at Urbana-Champaign. Email: {tavildar,pramodv}@uiuc.edu.

***Abstract*— We consider a problem in which two encoders observe different components of a memoryless, Gaussian, vector-valued source. The encoders separately communicate with a decoder, which attempts to reproduce the vector-valued source subject to constraints on the expected squared error of each component. We complete the determination of the rate region for this problem by determining the minimum sum rate needed to meet a pair of target distortions. The proof involves coupling the problem to a quadratic Gaussian "CEO problem."**

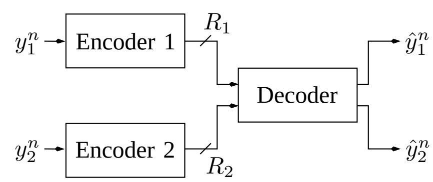

Fig. 1. The two-terminal source-coding problem.

I. Introduction

This paper addresses the quadratic Gaussian version of the two-terminal source-coding problem, the setup for which is depicted in Fig. 1. Two encoders observe different components of a memoryless, Gaussian, vector-valued source. We assume that these components are correlated in general. The encoders, without cooperating, send messages to a single decoder over rate-constrained, noiseless channels. The decoder attempts to reproduce both components, subject to separate constraints on the expected squared error of its estimates. We seek to determine the set of rate pairs (R_1, R_2) that allow the decoder to meet a given pair of target distortions. We call this set the *rate region*. Of course, this problem can be easily formulated for general sources and distortion measures. Our focus on the quadratic Gaussian case is motivated both by its fundamental nature and by its importance in applications. This problem is naturally viewed as a lossy version of Slepian and Wolf's problem [1]. The main result of this paper is an explicit characterization of the rate region as a function of the target distortions.

This problem was studied as early as 1978 [2], [3], when it was recognized that an inner bound could be obtained by combining vector quantization with the binning method of Cover [5]. More recently, Oohama [4] determined the rate region for the problem in which only one of the two distortion constraints is present. By interpreting this problem as a relaxation of the original problem, he obtained an outer bound on the rate region of the latter. He showed that this outer bound, when combined with the inner bound just mentioned, determines a portion of the boundary of the rate region. As a result of his work, showing that the inner bound is tight in the sum rate suffices to complete the characterization of the rate region. This is the contribution of the current paper.

Our approach is to lower bound the sum rate of a given code in two ways. The first way amounts to considering the rate required by a hypothetical centralized encoder that achieves the same error covariance matrix as the code. The second way is to establish a connection between this problem and the quadratic Gaussian "CEO problem," and then invoke existing results characterizing the sum rate of the latter. For some codes, the cooperative bound is tighter. For others, the CEO-based bound may be tighter. Taking the maximum of the two yields the desired lower bound.

The next section contains a precise formulation of the problem and the statement of our main result, Theorem 1. We provide some preliminaries and the necessary background on the CEO problem in Section III. We prove our main result in Section IV. Section V contains some concluding remarks.

We use the following notation. Boldface, lower case letters ($\boldsymbol{\mu}$) denote vectors over space while letters with a superscript n (y^n) denote vectors over time. Boldface, upper case letters ($\boldsymbol{D}$) denote matrices. Lightface letters (ρ, R) denote scalars. Whether a variable is deterministic or random should be clear from the context.

II. Problem Formulation and Main Result

Let $\{(y_1^n(i), y_2^n(i))\}_{i=1}^n$ be a sequence of independent and identically distributed (i.i.d.) Gaussian zero-mean random vectors. Let

$$\boldsymbol{K}_y = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$$

denote the covariance matrix of $(y_1^n(1), y_2^n(1))$. We use $y_1^n(j:k)$ to denote $\{y_1^n(i)\}_{i=j}^k$, $\mathbf{y}^n(i)$ to denote $(y_1^n(i), y_2^n(i))$, and $\mathbf{y}^n(j:k)$ to denote $\{(y_1^n(i), y_2^n(i))\}_{i=j}^k$. Analogous notation will be used for other random vectors.

The first encoder observes y_1^n, then sends a message to the decoder using the map

$$f_1^{(n)} : \mathbb{R}^n \mapsto \left\{1, \dots, M_1^{(n)}\right\}.$$

The second encoder operates analogously. The decoder uses the received messages to estimate both y_1^n and y_2^n according to the maps

$$\varphi_j^{(n)} : \left\{1, \dots, M_1^{(n)}\right\} \times \left\{1, \dots, M_2^{(n)}\right\} \mapsto \mathbb{R}^n \quad j = 1, 2.$$

Definition 1: A rate-distortion vector (R_1, R_2, d_1, d_2) is ***achievable*** if there exists a block length n, encoders $(f_1^{(n)}, f_2^{(n)})$, and a decoder $(\varphi_1^{(n)}, \varphi_2^{(n)})$ such that[1]

$$\begin{aligned} R_j &\geq \frac{1}{n} \log M_j^{(n)} \text{ for all } j \text{ in } \{1,2\}, \text{ and} \\ d_j &\geq \frac{1}{n} \sum_{i=1}^{n} E\left[\left(y_j^n(i) - \hat{y}_j^n(i)\right)^2\right] \text{ for all } j \text{ in } \{1,2\}, \end{aligned} \tag{1}$$

where

$$\hat{y}_j^n = \varphi_j^{(n)}\left(f_1^{(n)}(y_1^n), f_2^{(n)}(y_2^n)\right) \text{ for } j \text{ in } \{1,2\}.$$

Let $\mathcal{RD}^\star$ denote the set of achievable rate-distortion vectors, and let

$$\mathcal{R}^\star(d_1, d_2) = \left\{(R_1, R_2) : (R_1, R_2, d_1, d_2) \in \overline{\mathcal{RD}^\star}\right\},$$

where $\overline{\mathcal{RD}^\star}$ denotes the closure of $\mathcal{RD}^\star$. We call $\mathcal{R}^\star(\cdot, \cdot)$ the *rate region* for the problem. The ***(minimum) sum rate*** for a given distortion pair (d_1, d_2) is defined to be

$$\inf\{R_1 + R_2 : (R_1, R_2) \in \mathcal{R}^\star(d_1, d_2)\}.$$

We note that there is no loss of generality in assuming that $E[y_1^n(1)] = E[y_2^n(1)] = 1$, since the observations and the estimates can be scaled to reduce the general case to this one. By similar reasoning, we may assume that $\rho \geq 0$, i.e., that the observations of the two encoders are nonnegatively correlated. Since the two extreme cases $\rho = 0$ and $\rho = 1$ can be handled using classical techniques, we shall assume throughout the remainder of the paper that $0 < \rho < 1$.

We now define three sets that will be used to describe the rate region for this problem. Let

$$\mathcal{R}_1^\star(d_1) = \Bigg\{(R_1, R_2) : R_1 \geq \frac{1}{2} \log^+ \left[\frac{1}{d_1}\left(1 - \rho^2 + \rho^2 2^{-2R_2}\right)\right]\Bigg\},$$

where $\log^+ x = \max(\log x, 0)$. Likewise, let

$$\mathcal{R}_2^\star(d_2) = \Bigg\{(R_1, R_2) : R_2 \geq \frac{1}{2} \log^+ \left[\frac{1}{d_2}\left(1 - \rho^2 + \rho^2 2^{-2R_1}\right)\right]\Bigg\}.$$

[1]All logarithms in this paper are base two.

Finally, let

$$\mathcal{R}_{12}^\star(d_1, d_2) = \Bigg\{(R_1, R_2) : R_1 + R_2 \geq \frac{1}{2} \log^+ \left[\frac{(1 - \rho^2)\, \beta(d_1, d_2)}{2 d_1 d_2}\right]\Bigg\},$$

where

$$\beta(d_1, d_2) = 1 + \sqrt{1 + \frac{4\rho^2 d_1 d_2}{(1 - \rho^2)^2}}.$$

Theorem 1: For the Gaussian two-terminal source-coding problem,

$$\mathcal{R}^\star(d_1, d_2) = \mathcal{R}_1^\star(d_1) \cap \mathcal{R}_2^\star(d_2) \cap \mathcal{R}_{12}^\star(d_1, d_2).$$

The proof is deferred to Section IV. In the meantime, we review the quadratic Gaussian CEO problem and state some ancillary results.

III. Background and Preliminaries

There is a natural coding method for this problem. Each encoder first vector quantizes its observation as in single-user rate-distortion theory. The resulting digital signals are then communicated losslessly to the decoder via Slepian-Wolf coding [1]. The decoder uses the quantizations to estimate the observations. Using this method, one can show one half of Theorem 1 [2], [3], [4],

$$\mathcal{R}_1^\star(d_1) \cap \mathcal{R}_2^\star(d_2) \cap \mathcal{R}_{12}^\star(d_1, d_2) \subset \mathcal{R}^\star(d_1, d_2). \tag{2}$$

Oohama [4] found the rate region when only one of the two distortion constraints is present

$$\begin{aligned} \mathcal{R}^\star(d_1, 1) &= \mathcal{R}_1^\star(d_1) \\ \mathcal{R}^\star(1, d_2) &= \mathcal{R}_2^\star(d_2). \end{aligned}$$

As a consequence, it holds

$$\mathcal{R}^\star(d_1, d_2) \subset \mathcal{R}_1^\star(d_1) \cap \mathcal{R}_2^\star(d_2). \tag{3}$$

These results already determine the rate region in a special case. To describe this case, let $\mathcal{D}_G$ be the set of all 2×2 matrices $\boldsymbol{D}$ such that

$$\boldsymbol{D}^{-1} = \boldsymbol{K}_y^{-1} + \boldsymbol{\Lambda}^{-1} \tag{4}$$

for some diagonal and positive definite matrix $\boldsymbol{\Lambda}$. We call $\mathcal{D}_G$ the set of *distributed Gaussian distortion matrices*, since it equals the set of matrices $\boldsymbol{D}$ such that there exist random variables u_1 and u_2 satisfying the following four properties:

- (*i*) (y_1, y_2, u_1, u_2) are jointly Gaussian,
- (*ii*) $u_1 \leftrightarrow y_1 \leftrightarrow y_2 \leftrightarrow u_2$, meaning that u_1, y_1, y_2, and u_2 form a Markov chain in this order,
- (*iii*) $0 < \mathrm{Var}(y_j | u_j) < 1$ for all $j \in \{1, 2\}$, and
- (*iv*) the covariance matrix of (y_1, y_2) given (u_1, u_2), $\mathrm{Cov}(\mathbf{y}|\mathbf{u})$, is equal to $\boldsymbol{D}$.

Let $\mathrm{diag}(\mathcal{D}_G)$ denote the set of distortion pairs (d_1, d_2) such that there exists a distributed Gaussian distortion matrix $\boldsymbol{D}$ with top-left element d_1 and bottom-right element d_2. If

(d_1, d_2) is not in $\mathrm{diag}(\mathcal{D}_G)$, then the rate region can be determined using existing results.

Lemma 1: If (d_1, d_2) is not in $\mathrm{diag}(\mathcal{D}_G)$, then

$$\mathcal{R}^\star(d_1, d_2) = \mathcal{R}_1^\star(d_1) \cap \mathcal{R}_2^\star(d_2) \cap \mathcal{R}_{12}^\star(d_1, d_2).$$

The proof requires only an elementary calculation and is omitted. In light of this lemma and Eqs. (2) and (3), it suffices to show that when (d_1, d_2) is in $\mathrm{diag}(\mathcal{D}_G)$,

$$\mathcal{R}^\star(d_1, d_2) \subset \mathcal{R}_{12}^\star(d_1, d_2).$$

This is shown in the next section.

Our proof uses a characterization of the sum rate for the quadratic Gaussian CEO problem. In the two-encoder version of this problem, encoders 1 and 2 observe y_1 and y_2, respectively, and then communicate with a single decoder as in the original problem. But now y_1 and y_2 are of the form

$$y_j = a_j x + n_j \quad j \text{ in } \{1, 2\},$$

where x, n_1, and n_2 are independent, Gaussian random variables, and the decoder estimates x instead of $\mathbf{y}$. The distortion measure is again the expected squared error. The rate region for this problem was recently found independently by Oohama [6] and Prabhakaran, Tse, and Ramchandran [7][2]. The sum rate can be expressed as the solution to an optimization problem over Gaussian auxiliary random variables u_1 and u_2

$$\inf \Big\{ I(y_1, y_2; u_1, u_2) : (x, \mathbf{y}, \mathbf{u}) \text{ are jointly Gaussian}, \\ u_1 \leftrightarrow y_1 \leftrightarrow y_2 \leftrightarrow u_2, \\ x \leftrightarrow \mathbf{y} \leftrightarrow \mathbf{u}, \text{ and} \\ E\left[(x - E[x|\mathbf{u}])^2\right] \le d \Big\}, \tag{5}$$

where d is the allowable distortion.

For our present purpose, we will find it more convenient to consider the related problem in which the decoder attempts to estimate $\boldsymbol{\mu}^T \mathbf{y}$ for some given vector $\boldsymbol{\mu}$. We call this problem the ***μ*-*sum problem***. For some values of $\boldsymbol{\mu}$, the $\boldsymbol{\mu}$-sum problem can be easily solved by coupling it to a CEO problem.

Lemma 2: The sum rate for the $\boldsymbol{\mu}$-sum problem with $\mu_1 \cdot \mu_2 > 0$ and allowable distortion d is given by

$$\inf \left\{ \frac{1}{2} \log \frac{|\boldsymbol{K}_y|}{|\boldsymbol{D}|} : \boldsymbol{D} \in \mathcal{D}_G \text{ and } \boldsymbol{\mu}^T \boldsymbol{D} \boldsymbol{\mu} \le d \right\}. \tag{6}$$

The proofs of this lemma and the next one are omitted due to space constraints. Both can be found in the full version of the paper [8]. Here we consider some properties of $\mathcal{D}_G$ and the optimization problem (6). Recall that $\boldsymbol{D}$ is in $\mathcal{D}_G$ if there exists a diagonal and positive definite matrix $\boldsymbol{\Lambda}$ such that

$$\boldsymbol{D}^{-1} = \boldsymbol{K}_y^{-1} + \boldsymbol{\Lambda}^{-1}. \tag{7}$$

This formula provides a convenient way of evaluating the off-diagonal element of $\boldsymbol{D}$ in terms of its diagonal elements and ρ. Let us write

$$\boldsymbol{D} = \begin{bmatrix} d_1 & \theta\sqrt{d_1 d_2} \\ \theta\sqrt{d_1 d_2} & d_2 \end{bmatrix},$$

[2] In fact, both works solved the problem for an arbitrary number of encoders, but this generality is not needed here.

where $\theta \in (-1, 1)$. Equating the off-diagonal elements in (7) gives

$$\frac{\theta}{(1-\theta^2)\sqrt{d_1 d_2}} = \frac{\rho}{1-\rho^2}.$$

This equation has several uses:

1) It allows us to immediately conclude that θ must be positive.
2) It shows that there is a unique $\boldsymbol{D}$ in $\mathcal{D}_G$ with top-left element d_1 and bottom-right element d_2, since this quadratic equation in θ only has one solution in $(-1, 1)$.
3) By solving for θ and taking the solution in $(-1, 1)$, it yields a formula for the determinant of $\boldsymbol{D}$,

$$|\boldsymbol{D}| = \frac{2 d_1 d_2}{\beta(d_1, d_2)}, \tag{8}$$

where $\beta(\cdot, \cdot)$ was defined in the last section.

It can happen that the infimum in (6) is not achieved by any $\boldsymbol{D}$ in $\mathcal{D}_G$. But it turns out that every $\boldsymbol{D}$ in $\mathcal{D}_G$ solves a $\boldsymbol{\mu}$-sum problem for some $\boldsymbol{\mu}$ with $\mu_1 \cdot \mu_2 > 0$. This fact will be used in the proof of our main result.

Lemma 3: For any $\boldsymbol{D}^*$ in $\mathcal{D}_G$, there is a vector $\boldsymbol{\mu}$ with $\mu_1 \cdot \mu_2 > 0$ such that $\boldsymbol{D}^*$ achieves the sum rate for the $\boldsymbol{\mu}$-sum problem, i.e.,

$$\frac{1}{2} \log \frac{|\boldsymbol{K}_y|}{|\boldsymbol{D}^*|} = \\ \inf \left\{ \frac{1}{2} \log \frac{|\boldsymbol{K}_y|}{|\boldsymbol{D}|} : \boldsymbol{D} \in \mathcal{D}_G \text{ and } \boldsymbol{\mu}^T \boldsymbol{D} \boldsymbol{\mu} \le \boldsymbol{\mu}^T \boldsymbol{D}^* \boldsymbol{\mu} \right\}.$$

The proof of this lemma in the full version of the paper [8] shows that if $\boldsymbol{D}^*$ has equal diagonal elements, then $\boldsymbol{D}^*$ solves the $\boldsymbol{\mu}$-CEO problem with $\boldsymbol{\mu}^T$ equal to $[1 \quad 1]$. Thus, we can explicitly identify the vector returned by Lemma 3 in this case. This fact makes the proofs that follow simpler and more concrete when the two distortion constraints, d_1 and d_2, are equal. As such, the reader is encouraged to keep this case in mind as we turn to the proof of the main result.

IV. Proof of the Main Result

Recall that we may restrict attention to the case in which (d_1, d_2) is in $\mathrm{diag}(\mathcal{D}_G)$. Let us now fix one such point; we will suppress dependence on (d_1, d_2) in what follows. Let $\boldsymbol{D}^*$ denote the element of $\mathcal{D}_G$ whose top-left and bottom-right elements are d_1 and d_2, respectively.

Definition 2: For $\theta \in (-1, 1)$, let

$$\boldsymbol{D}_\theta = \begin{bmatrix} d_1 & \theta\sqrt{d_1 d_2} \\ \theta\sqrt{d_1 d_2} & d_2 \end{bmatrix},$$

and define

$$R_{\mathrm{coop}}(\theta) = \frac{1}{2} \log^+ \frac{|\boldsymbol{K}_y|}{|\boldsymbol{D}_\theta|} = \frac{1}{2} \log^+ \frac{1-\rho^2}{(1-\theta^2) d_1 d_2}.$$

Let $\boldsymbol{\mu}$ denote the vector supplied by Lemma 3, that is, a vector with $\mu_1 \cdot \mu_2 > 0$ such that

$$\frac{1}{2} \log \frac{|\boldsymbol{K}_y|}{|\boldsymbol{D}^*|} = \\ \inf \left\{ \frac{1}{2} \log \frac{|\boldsymbol{K}_y|}{|\boldsymbol{D}|} : \boldsymbol{D} \in \mathcal{D}_G \text{ and } \boldsymbol{\mu}^T \boldsymbol{D} \boldsymbol{\mu} \le \boldsymbol{\mu}^T \boldsymbol{D}^* \boldsymbol{\mu} \right\}.$$

Then define

$$R_{\mathrm{sum}}(\theta) = \inf\left\{\frac{1}{2}\log\frac{|\boldsymbol{K}_y|}{|\boldsymbol{D}|} : \boldsymbol{D}\in\mathcal{D}_G \text{ and } \boldsymbol{\mu}^T\boldsymbol{D}\boldsymbol{\mu} \le \boldsymbol{\mu}^T\boldsymbol{D}_\theta\boldsymbol{\mu}\right\}.$$

The next lemma is central to the proof of our main result.

Lemma 4: If (R_1, R_2, d_1, d_2) is achievable, then

$$R_1 + R_2 \ge \inf_{\theta\in(-1,1)} \max\left(R_{\mathrm{coop}}(\theta), R_{\mathrm{sum}}(\theta)\right). \quad (9)$$

Proof: The hypothesis implies that there exists a code $(f_1^{(n)}, f_2^{(n)}, \varphi_1^{(n)}, \varphi_2^{(n)})$ satisfying (1). Then

$$\begin{aligned} n(R_1+R_2) &\ge h\left(f_1^{(n)}(y_1^n), f_2^{(n)}(y_2^n)\right) \\ &= I\left(\mathbf{y}^n; f_1^{(n)}(y_1^n), f_2^{(n)}(y_2^n)\right) \\ &= h(\mathbf{y}^n) - h\left(\mathbf{y}^n \middle| f_1^{(n)}(y_1^n), f_2^{(n)}(y_2^n)\right), \quad (10) \end{aligned}$$

where $h(\cdot)$ denotes differential entropy. But

$$h(\mathbf{y}^n) = \frac{n}{2}\log\left[(2\pi e)^2|\boldsymbol{K}_y|\right] \quad (11)$$

[9, Theorem 9.4.1], and

$$\begin{aligned} & h\left(\mathbf{y}^n \middle| f_1^{(n)}(y_1^n), f_2^{(n)}(y_2^n)\right) \\ &= \sum_{i=1}^n h\left(\mathbf{y}^n(i) \middle| f_1^{(n)}(y_1^n), f_2^{(n)}(y_2^n), \mathbf{y}^n(1:i-1)\right) \\ &\le \sum_{i=1}^n h\left(\mathbf{y}^n(i) - \hat{\mathbf{y}}^n(i) \middle| f_1^{(n)}(y_1^n), f_2^{(n)}(y_2^n)\right) \\ &\le \sum_{i=1}^n h\left(\mathbf{y}^n(i) - \hat{\mathbf{y}}^n(i)\right), \end{aligned}$$

since conditioning reduces entropy. Now let $\tilde{\boldsymbol{D}}_i$ denote the distortion matrix of $\mathbf{y}^n(i)$,

$$\tilde{\boldsymbol{D}}_i = E\left[\left(\mathbf{y}^n(i) - \hat{\mathbf{y}}^n(i)\right)\left(\mathbf{y}^n(i) - \hat{\mathbf{y}}^n(i)\right)^T\right],$$

and let

$$\tilde{\boldsymbol{D}} = \frac{1}{n}\sum_{i=1}^n \tilde{\boldsymbol{D}}_i$$

denote the average distortion matrix of the code. We may assume that $\varphi_1^{(n)}$ and $\varphi_2^{(n)}$ are MMSE estimators, in which case Theorem 9.6.5 in Cover and Thomas [9] implies that

$$h\left(\mathbf{y}^n(i) - \hat{\mathbf{y}}^n(i)\right) \le \frac{1}{2}\log\left[(2\pi e)^2|\tilde{\boldsymbol{D}}_i|\right].$$

Applying the concavity of log-det [9, Theorem 16.8.1], we have

$$\begin{aligned} \frac{1}{n} h\left(\mathbf{y}^n \middle| f_1^{(n)}(y_1^n), f_2^{(n)}(y_2^n)\right) &\le \frac{1}{n}\sum_{i=1}^n \frac{1}{2}\log\left[(2\pi e)^2|\tilde{\boldsymbol{D}}_i|\right] \\ &\le \frac{1}{2}\log\left[(2\pi e)^2|\tilde{\boldsymbol{D}}|\right]. \end{aligned}$$

Now

$$h\left(\mathbf{y}^n \middle| f_1^{(n)}(y_1^n), f_2^{(n)}(y_2^n)\right) > -\infty$$

by (10). Thus $\tilde{\boldsymbol{D}}$ must be nonsingular and hence positive definite. Let us write it as

$$\tilde{\boldsymbol{D}} = \begin{bmatrix} \tilde{d}_1 & \tilde{\theta}\sqrt{\tilde{d}_1\tilde{d}_2} \\ \tilde{\theta}\sqrt{\tilde{d}_1\tilde{d}_2} & \tilde{d}_2 \end{bmatrix},$$

where $\tilde{d}_1 \le d_1$, $\tilde{d}_2 \le d_2$, and $\tilde{\theta}$ is in $(-1,1)$. Now define

$$\theta = \frac{\tilde{\theta}\sqrt{\tilde{d}_1\tilde{d}_2}}{\sqrt{d_1 d_2}}.$$

Then $\tilde{\boldsymbol{D}} \preceq \boldsymbol{D}_\theta$, meaning that $\boldsymbol{D}_\theta - \tilde{\boldsymbol{D}}$ is positive semidefinite. In particular, $|\tilde{\boldsymbol{D}}| \le |\boldsymbol{D}_\theta|$ [10, Corollary 7.7.4]. This implies

$$\frac{1}{n} h\left(\mathbf{y}^n \middle| f_1^{(n)}(y_1^n), f_2^{(n)}(y_2^n)\right) \le \frac{1}{2}\log\left[(2\pi e)^2|\boldsymbol{D}_\theta|\right]. \quad (12)$$

Combining Eqs. (10), (11), and (12) gives

$$R_1 + R_2 \ge R_{\mathrm{coop}}(\theta). \quad (13)$$

Next observe that

$$E\left[\left(\boldsymbol{\mu}^T\mathbf{y}^n(i) - \boldsymbol{\mu}^T\hat{\mathbf{y}}^n(i)\right)^2\right] = \boldsymbol{\mu}^T\tilde{\boldsymbol{D}}_i\boldsymbol{\mu}.$$

In particular,

$$\frac{1}{n}\sum_{i=1}^n E\left[\left(\boldsymbol{\mu}^T\mathbf{y}^n(i) - \boldsymbol{\mu}^T\hat{\mathbf{y}}^n(i)\right)^2\right] \le \boldsymbol{\mu}^T\boldsymbol{D}_\theta\boldsymbol{\mu},$$

i.e., this code achieves distortion at most $\boldsymbol{\mu}^T\boldsymbol{D}_\theta\boldsymbol{\mu}$ for the $\boldsymbol{\mu}$-sum problem. Lemma 2 then implies that

$$R_1 + R_2 \ge R_{\mathrm{sum}}(\theta).$$

Combining this with (13) gives

$$R_1 + R_2 \ge \max(R_{\mathrm{coop}}(\theta), R_{\mathrm{sum}}(\theta)).$$

The conclusion follows by taking the infimum over θ in $(-1,1)$. □

The next step is to evaluate the infimum in (9). The two bounds $R_{\mathrm{coop}}(\cdot)$ and $R_{\mathrm{sum}}(\cdot)$ are shown in Fig. 2 for the case $\rho = 0.5$ and $d_1 = d_2 = 0.5$. We will show that these two functions always intersect at the correlation coefficient of $\boldsymbol{D}^*$, and at this point, they equal the min-max.

Lemma 5: It holds,

$$\inf_{\theta\in(-1,1)} \max\left(R_{\mathrm{coop}}(\theta), R_{\mathrm{sum}}(\theta)\right) = \frac{1}{2}\log^+\left[\frac{(1-\rho^2)\,\beta(d_1,d_2)}{2d_1d_2}\right].$$

Proof: Let us write $\boldsymbol{D}^*$, the matrix in $\mathcal{D}_G$ with diagonal entries (d_1, d_2), as

$$\boldsymbol{D}^* = \begin{bmatrix} d_1 & \theta^*\sqrt{d_1d_2} \\ \theta^*\sqrt{d_1d_2} & d_2 \end{bmatrix}.$$

Then observe that since $\theta^* > 0$, if $\theta > \theta^*$, we have

$$\begin{aligned} \max(R_{\mathrm{coop}}(\theta), R_{\mathrm{sum}}(\theta)) &\ge R_{\mathrm{coop}}(\theta) \\ &\ge R_{\mathrm{coop}}(\theta^*) = \frac{1}{2}\log\frac{|\boldsymbol{K}_y|}{|\boldsymbol{D}^*|}. \end{aligned}$$

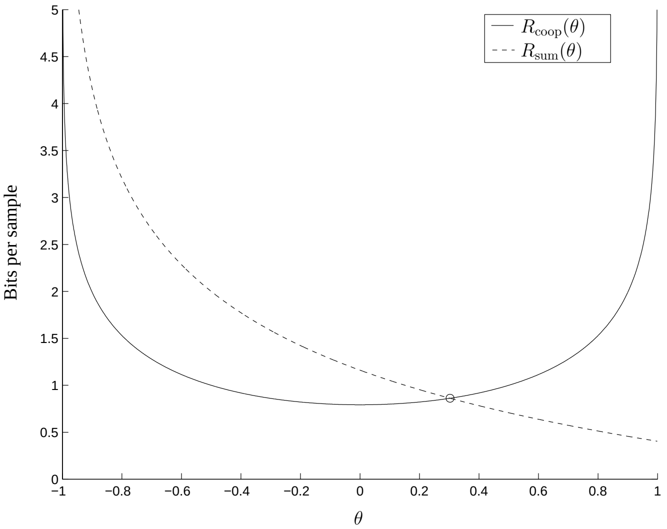

Fig. 2. $R_{\mathrm{coop}}(\cdot)$ and $R_{\mathrm{sum}}(\cdot)$ for the case $\rho = 0.5$ and $d_1 = d_2 = 0.5$. The $\boldsymbol{\mu}$ vector in the definition of $R_{\mathrm{sum}}(\cdot)$ is taken to be $[1 \ \ 1]^T$. The point at which the two functions intersect is the min-max and equals the sum rate.

On the other hand, if $\theta < \theta^*$, then since $R_{\mathrm{sum}}(\cdot)$ is nonincreasing,

$$\max(R_{\mathrm{coop}}(\theta), R_{\mathrm{sum}}(\theta)) \geq R_{\mathrm{sum}}(\theta) \geq R_{\mathrm{sum}}(\theta^*) = \frac{1}{2}\log\frac{|\boldsymbol{K}_y|}{|\boldsymbol{D}^*|},$$

where we have used the fact that $\boldsymbol{\mu}$ in the definition of $R_{\mathrm{sum}}(\cdot)$ was chosen so that $\boldsymbol{D}^*$ achieves the sum rate for the $\boldsymbol{\mu}$-sum problem with distortion $\boldsymbol{\mu}^T\boldsymbol{D}^*\boldsymbol{\mu}$. It follows that

$$\inf_{\theta\in(-1,1)} \max(R_{\mathrm{coop}}(\theta), R_{\mathrm{sum}}(\theta)) = \frac{1}{2}\log\frac{|\boldsymbol{K}_y|}{|\boldsymbol{D}^*|}.$$

Lemma 2 implies that $R_{\mathrm{sum}}(\cdot)$ is nonnegative. Hence, so is the left-hand side. We can conclude the proof, therefore, by invoking the formula for the determinant of a distributed Gaussian distortion matrix, Eq. (8). □

Proof of Theorem 1. As discussed in Section III, it suffices to show that

$$\mathcal{R}^\star(d_1, d_2) \subset \mathcal{R}^\star_{12}(d_1, d_2)$$

when (d_1, d_2) is in $\mathrm{diag}(\mathcal{D}_G)$. In this case, Lemmas 4 and 5 imply that if the rate-distortion vector (R_1, R_2, d_1, d_2) is achievable, then

$$R_1 + R_2 \geq \frac{1}{2}\log^+\left[\frac{(1-\rho^2)\ \beta(d_1, d_2)}{2d_1 d_2}\right]. \qquad (14)$$

Since the right-hand side is continuous in (d_1, d_2), it follows that if the point (R_1, R_2, d_1, d_2) is in $\overline{\mathcal{RD}^\star}$, then (14) again holds. This implies the desired conclusion. □

V. Discussion

A consequence of our result is that single-user vector quantization followed by Slepian-Wolf coding [2], [3], [4] achieves the entire rate region for this problem. Recent work has shown this technique to be optimum for several network source-coding problems. For instance, it has been shown to achieve the entire rate region of the quadratic Gaussian CEO problem [6], [7] and the sum rate of the binary erasure CEO problem [11].

Our result relies on the solution to the $\boldsymbol{\mu}$-sum problem given in Lemma 2. The proof of Lemma 2, which is given in the full version of the paper [8], is noteworthy in that it involves adding a new component x to the source (y_1, y_2). Unlike more typical auxiliary random variables, x does not represent a component of the code. Rather, it is used to aid the analysis by inducing conditional independence among the messages sent by the two encoders. This technique of augmenting the source to induce conditional independence has proven useful in other contexts as well. Ozarow [12] used it to prove the converse for the Gaussian two-descriptions problem. Wang and Viswanath [13] used it to determine the sum rate for the Gaussian vector multiple-descriptions problem with individual and central decoders. Finally, Wagner and Anantharam [11] used it to prove an outer bound for the general multiterminal source-coding problem.

Acknowledgment

The authors would like to thank Venkat Anantharam and Jun Chen for helpful discussions.

References

[1] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. 19, no. 4, pp. 471–480, July 1973.

[2] T. Berger, "Multiterminal source coding," in *The Information Theory Approach to Communications*, ser. CISM Courses and Lectures, G. Longo, Ed. Springer-Verlag, 1978, vol. 229, pp. 171–231.

[3] S.-Y. Tung, "Multiterminal source coding," Ph.D. dissertation, School of Electrical Engineering, Cornell University, Ithaca, NY, May 1978.

[4] Y. Oohama, "Gaussian multiterminal source coding," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1912–1923, Nov. 1997.

[5] T. M. Cover, "A proof of the data compression theorem of Slepian and Wolf for ergodic sources," *IEEE Trans. Inf. Theory*, vol. 21, no. 2, pp. 226–228, Mar. 1975.

[6] Y. Oohama, "Rate-distortion theory for Gaussian multiterminal source coding systems with several side informations at the decoder," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2577–2593, July 2005.

[7] V. Prabhakaran, D. Tse, and K. Ramchandran, "Rate region of the quadratic Gaussian CEO problem," in *IEEE Int. Symp. Inf. Theor. Proc.*, 2004, p. 117.

[8] A. B. Wagner, S. Tavildar, and P. Viswanath, "The rate region of the quadratic Gaussian two-terminal source-coding problem," `arXiv:cs.IT/0510095`.

[9] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: John Wiley & Sons, 1991.

[10] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.

[11] A. B. Wagner and V. Anantharam, "An infeasibility result for the multiterminal source-coding problem," `arXiv:cs.IT/0511103`.

[12] L. Ozarow, "On a source-coding problem with two channels and three receivers," *Bell Syst. Tech. J.*, vol. 59, no. 10, pp. 1909–1921, Dec. 1980.

[13] H. Wang and P. Viswanath, "Vector Gaussian multiple description with individual and central receivers," `arXiv:cs.IT/0510078`.